SYSTEM ARCHITECTURE FOR REMOTE ACCESS AND CONTROL OF ENVIRONMENTAL MANAGEMENT

Related Applications

The subject matter of U.S. Patent Application entitled "Method of Remote Access and Control of Environmental Management," filed on October 1, 1997, Application No. 08/942-215 and having attorney Docket No. MNFRAME.002A2 is related to this application.

Priority Claim

The benefit under 35 U.S.C. § 119(e) of the following U.S. provisional application(s) is hereby claimed:

Title	Application No.	Filing Date
"Remote Access and Control of Environmental Management System"	60/046,397	May 13, 1997
"Hardware and Software Architecture for Inter-Connecting an Environmental Management System with a Remote Interface"	60/047,016	May 13, 1997
"Self Management Protocol for a Fly-By-Wire Service Processor"	60/046,416	May 13, 1997

Appendices

Appendix A, which forms a part of this disclosure, is a list of commonly owned copending U.S. patent applications. Each one of the applications listed in Appendix A is hereby incorporated herein in its entirety by reference thereto.

Appendix B, which forms part of this disclosure, is a copy of the U.S. provisional patent application filed May 13, 1997, entitled "Remote Access and Control of Environmental Management System" and assigned Application No. 60/046,397. Page 1, line 6 of the provisional application has been changed from the

original to positively recite that the entire provisional application, including the attached documents, forms part of this disclosure.

Copyright Rights

A portion of the disclosure of this patent document contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the Patent and Trademark Office patent files or records, but otherwise reserves all copyright rights whatsoever.

Background of the Invention

Field of the Invention

The invention relates to fault tolerant computer systems. More specifically, the invention is directed to a system for providing remote access and control of server environmental management.

Description of the Related Technology

As enterprise-class servers become more powerful and more capable, they are also becoming increasingly sophisticated and complex. For many companies, these changes lead to concerns over server reliability and manageability, particularly in light of the increasingly critical role of server-based applications. While in the past many systems administrators were comfortable with all of the various components that made up a standards-based network server, today's generation of servers can appear as an incomprehensible, unmanageable black box. Without visibility into the underlying behavior of the system, the administrator must "fly blind." Too often the only indicators the network manager has on the relative health of a particular server is whether or not it is running.

It is well-acknowledged that there is a lack of reliability and availability of most standards-based servers. Server downtime, resulting either from hardware or software faults or from regular maintenance, continues to be a significant problem. By one estimate, the cost of downtime in mission critical environments has risen to an annual total of $4.0 billion for U.S. businesses, with the average downtime event

resulting in a $140 thousand loss in the retail industry and a $450 thousand loss in the securities industry. It has been reported that companies lose as much as $250 thousand in employee productivity for every 1% of computer downtime. With emerging Internet, intranet and collaborative applications taking on more essential business roles every day, the cost of network server downtime will continue to spiral upward.

While hardware fault tolerance is an important element of an overall high availability architecture, it is only one piece of the puzzle. Studies show that a significant percentage of network server downtime is caused by transient faults in the I/O subsystem. These faults may be due, for example, to the device driver, the adapter card firmware, or hardware which does not properly handle concurrent errors, and often causes servers to crash or hang. The result is hours of downtime per failure, while a system administrator discovers the failure takes some action, and manually reboots the server. In many cases, data volumes on hard disk drives become corrupt and must be repaired when the volume is mounted. A dismount-and-mount cycle may result from the lack of "hot pluggability" in current standards-based servers. Diagnosing intermittent errors can be a frustrating and time-consuming process. For a system to deliver consistently high availability, it must be resilient to these types of faults. Accurate and available information about such faults is central to diagnosing the underlying problems and taking corrective action.

Modern fault tolerant systems have the functionality to provide the ambient temperature of a storage device enclosure and the operational status of other components such as the cooling fans and power supply. However, a limitation of these server systems is that they do not contain self-managing processes to correct malfunctions. Also, if a malfunction occurs in a typical server, it relies on the operating system software to report, record and manage recovery of the fault. However, many types of faults will prevent such software from carrying out these tasks. For example, a disk drive failure can prevent recording of the fault in a log file on that disk drive. If the system error caused the system to power down, then the system administrator would never know the source of the error.

Traditional systems are lacking in detail and sophistication when notifying system administrators of system malfunctions. System administrators are in need of a graphical user interface for monitoring the health of a network of servers. Administrators need a simple point-and-click interface to evaluate the health of each server in the network. In addition, existing fault tolerant servers rely upon operating system maintained logs for error recording. These systems are not capable of maintaining information when the operating system is inoperable due to a system malfunction. Existing systems do not have a system log for maintaining information when the main computational processors are inoperable.

Another limitation of the typical fault tolerant system is that the control logic for the diagnostic system is associated with a particular processor. Thus, if the environmental control processor malfunctioned, then all diagnostic activity on the computer would cease. In traditional systems, if a controller dedicated to the fan system failed, then all fan activity could cease resulting in overheating and ultimate failure of the server. What is desired is a way to obtain diagnostic information when the server OS is not operational or even when main power to the server is down.

Existing fault tolerant systems also lack the power to remotely control a particular server, such as powering up and down, resetting, retrieving or updating system status, displaying flight recorder and so forth. Such control of the server is desired even when the server power is down. For example, if the operating system on the remote machine failed, then a system administrator would have to physically go to the remote machine to re-boot the malfunctioning machine before any system information could be obtained or diagnostics could be started.

Therefore, a need exists for improvements in server management which will result in greater reliability and dependability of operation. Server users are in need of a management system by which the users can accurately gauge the health of their system. Users need a high availability system that must not only be resilient to faults, but must allow for maintenance, modification, and growth--without downtime. System users must be able to replace failed components, and add new functionality, such as new network interfaces, disk interface cards and storage, without impacting existing users. As system demands grow, organizations must frequently expand, or

scale, their computing infrastructure, adding new processing power, memory, storage and I/O capacity. With demand for 24-hour access to critical, server-based information resources, planned system downtime for system service or expansion has become unacceptable.

Summary of the Invention

Embodiments of the inventive remote access system provides system administrators with new levels of client/server system availability and management. It gives system administrators and network managers a comprehensive view into the underlying health of the server--in real time, whether on-site or off-site. In the event of a failure, the invention enables the administrator to learn why the system failed, why the system was unable to boot, and to control certain functions of the server from a remote station.

One embodiment of the present invention is a system for external management of a server environment, the system comprising a first computer having a plurality of components; a second computer connected to the first computer; a plurality of microcontrollers that monitor components of the first computer; a bus that interconnects the microcontrollers; and a remote interface circuit that interfaces the microcontrollers with the second computer thereby facilitating the external management of the first computer components.

Another embodiment of the present invention is a system for external management of a computer environment, the system comprising a first computer having a plurality of components; a second computer connected to the first computer; means for monitoring components of the first computer; and means for interfacing the monitoring means with the second computer to facilitate the external management of the first computer components.

Yet another embodiment of the present invention is a remote interface, comprising a microcontroller; a memory connected to the microcontroller; a first port capable of receiving and transmitting monitoring data; and a second port capable of receiving and transmitting monitoring data, wherein the second port includes connectivity to a microcontroller bus.

6

Brief Description of the Drawings

Figure 1 is a top level block diagram of microcontroller network components utilized by an embodiment of the present invention.

Figure 2 is a block diagram of the server portion of the microcontroller network shown in Figure 1.

Figure 3 is a block diagram of a remote interface board (RIB) that is part of the microcontroller network shown in Figures 1 and 2.

Figure 4 is a diagram of serial protocol message formats utilized by the RIB shown in Figure 3.

Figures 5a and 5b are a flowchart of a RIB microcontroller that is a part of the microcontroller network shown in Figures 1 and 2.

Figure 6 is a diagram of a modem dialing and answering state machine defined in Figure 5a.

Detailed Description of the Invention

The following detailed description presents a description of certain specific embodiments of the present invention. In this description, reference is made to the drawings wherein like parts are designated with like numerals throughout.

For convenience, the discussion of the invention is organized into the following principal sections: Introduction, Server System, Microcontroller Network, Remote Interface Board, Remote Interface Serial Protocol, and RIB Microcontroller Operation.

I. INTRODUCTION

The inventive computer server system and client computer includes a distributed hardware environment management system that is built as a small self-contained network of microcontrollers. Operating independently of the system processor and operating system software, embodiments of present invention uses separate processors for providing information and managing the hardware environment including fans, power supplies and temperature.

Initialization, modification and retrieval of system conditions are performed through utilization of a remote interface by issuing commands to the environmental processors. The system conditions may include system log size, presence of faults in the system log, serial number for each of the environmental processors, serial numbers for each power supply of the system, system identification, system log count, power settings and presence, canister presence, temperature, BUS/CORE speed ratio, fan speeds, settings for fan faults, LCD display, Non-Maskable Interrupt (NMI) request bits, CPU fault summary, FRU status, JTAG enable bit, system log information, remote access password, over-temperature fault, CPU error bits, CPU presence, CPU thermal fault bits, and remote port modem. The aforementioned list of capabilities provided by the present environmental system is not all-inclusive.

The server system and client computer provides mechanisms for the evaluation of the data that the system collects and methods for the diagnosis and repair of server problems in a manner that system errors can be effectively and efficiently managed. The time to evaluate and repair problems is minimized. The server system ensures that the system will not go down, so long as sufficient system resources are available to continue operation, but rather degrade gracefully until the faulty components can be replaced.

II. SERVER SYSTEM

Referring to Figure 1, a server system 100 with a remote client computer will be described. In a one embodiment, the server system hardware environment 100 may be built around a self-contained network of microcontrollers, such as, for example, a remote interface microcontroller on the remote interface board or circuit 104, a system interface microcontroller 106 and a system recorder microcontroller 110. This distributed service processor network 102 may operate as a fully self-contained subsystem within the server system 100, continuously monitoring and managing the physical environment of the machine (e.g., temperature, voltages, fan status). The microcontroller network 102 continues to operate and provides a system administrator with critical system information, regardless of the operational status of the server 100.

Information collected and analyzed by the microcontroller network 102 can be presented to a system administrator using either SNMP-based system management software (not shown), or using microcontroller network Recovery Manager software 130 through a local connection 121 or a dial-in connection 123. The system management software, which interfaces with the operating system (OS) 108 such as Microsoft Windows NT Version 4.0 or Novell Netware Version 4.11, for example, provides the ability to manage the specific characteristics of the server system, including Hot Plug Peripheral Component Interconnect (PCI), power and cooling status, as well as the ability to handle alerts associated with these features.

The microcontroller network Recovery Manager software 130 allows the system administrator to query the status of the server system 100 through the microcontroller network 102, even when the server is down. Using the microcontroller network remote management capability, a system administrator can use the Recovery Manager 130 to re-start a failed system through a modem connection 123. First, the administrator can remotely view the microcontroller network Flight Recorder, a feature that stores all system messages, status and error reports in a circular Non-Volatile Random Access Memory buffer (NVRAM) 112. Then, after determining the cause of the system problem, the administrator can use microcontroller network "fly by wire" capability to reset the system, as well as to power the system off or on. "Fly by wire" denotes that no switch, indicator or other control is directly connected to the function it monitors or controls, but instead, all the control and monitoring connections are made by the microcontroller network 102.

The remote interface board (RIB) 104 interfaces the server system 100 to an external client computer. The RIB 104 may be internal or external to an enclosure of the server 100. Furthermore, the RIB may be incorporated onto another circuit of the server, such as a system board 150 (Figure 2) or a backplane 152 of the server. The RIB 104 connects to either a local client computer 122 at the same location as the server 100 or to a remote client computer 124 through an optional switch 120. The client computer 122/124 may in one embodiment run either Microsoft Windows 95 or Windows NT Workstation version 4.0 operating system (OS) 132.

The client computer 122/124 could be another server, such as, for example, a backup server. The client computer 122/124 could also be a handheld computer such as, for example, a personal digital assistant (PDA). It is not necessary that Operating System software be running on the client computer 122/124. For example, the client computer 122/124 could be hard-wired for specific tasks, or could have special purpose embedded software.

The processor and RAM requirements of the client computer 122/124 are such as necessary by the OS 132. The serial port of the client computer 122/124 may utilize a type 16550A Universal Asynchronous Receiver Transmitter (UART). The switch 120 facilitates either the local connection 121 or the modem connection 123 at any one time, but allows both types of connections to be connected to the switch. In an another embodiment, either the local connection 121 or the modem connection 123 is connected directly to the RIB 104. The local connection 121 utilizes a readily available null-modem serial cable to connect to the local client computer. The modem connection may utilize a Hayes-compatible server modem 126 and a Hayes-compatible client modem 128. In one embodiment, a model V.34X 33.6K data/fax modem available from Zoom is utilized as the client modem and the server modem. In another embodiment, a Sportster 33.6K data/fax modem available from US Robotics is utilized as the client modem.

The steps of connecting the remote client computer 124 to the server 100 will now be briefly described. The remote interface 104 has a serial port connector 204 (Figure 3) that directly connects with a counterpart serial port connector of the external server modem 126 without the use of a cable. If desired, a serial cable could be used to interconnect the remote interface 104 and the server modem 126. The cable end of an AC to DC power adapter (not shown, for example a 120 Volt AC to 7.5 Volt DC, or a 220V, European or Japanese adapter) is then connected to the DC power connector J2 (220, Figure 3) of the remote interface, while the double-prong end is plugged into a 120 Volt AC wall outlet. One end of an RJ-45 parallel-wire data cable 103 is then plugged into an RJ-45 jack (226, Figure 3) on the remote interface 104, while the other end is plugged into a RJ-45 Recovery Manager jack on the server 100. The RJ-45 jack on the server then connects to the microcontroller

network 102. The server modem 126 is then connected to a communications network 127 using an appropriate connector. The communications network 127 may be a public switched telephone network, although other modem types and communication networks are envisioned. For example, if cable modems are used for the server modem 126 and client modem 128, the communications network can be a cable television network. As another example, satellite modulator/demodulators can be used in conjunction with a satellite network.

In another embodiment, the server modem to client modem connection may be implemented by an Internet connection utilizing the well known TCP/IP protocol. Any of several Internet access devices, such as modems or network interface cards, may be utilized. Thus, the communications network 127 may utilize either circuit or packet switching.

At the remote client computer 124, a serial cable (25-pin D-shell) 129 is used to interconnect the client modem 128 and the client computer 124. The client modem 128 is then connected to the communications network 127 using an appropriate connector. Each modem is then plugged into an appropriate power source for the modem, such as an AC outlet. At this time, the Recovery Manager software 130 is loaded into the client computer 124, if not already present, and activated.

The steps of connecting the local client computer 122 to the server 100 are similar, but modems are not necessary. The main difference is that the serial port connector of the remote interface 104 connects to a serial port of the local client computer 122 by the null-modem serial cable 121.

III. MICROCONTROLLER NETWORK

In one embodiment, the invention is implemented by a network of microcontrollers 102 (Figure 1). The microcontrollers may provide functionality for system control, diagnostic routines, self-maintenance control, and event logging processors. A further description of the microcontrollers and microcontroller network is provided in U.S. Patent Application No. 08/942-402 entitled "Diagnostic and Managing Distributed Processor System".

Referring to Figure 2, in one embodiment of the invention, the network of microcontrollers 102 includes ten processors. One of the purposes of the microcontroller network 102 is to transfer messages to the other components of the server system 100. The processors may include: a System Interface controller 106, a CPU A controller 166, a CPU B controller 168, a System Recorder 110, a Chassis controller 170, a Canister A controller 172, a Canister B controller 174, a Canister C controller 176, a Canister D controller 178 and a Remote Interface controller 200. The Remote Interface controller 200 is located on the RIB 104 (Figure 1) which is part of the server system 100, but may preferably be external to the server enclosure. The System Interface controller 106, the CPU A controller 166 and the CPU B controller 168 are located on thr system board 150 in the server 100. Also located on the system board are one or more central processing units (CPUs) or microprocessors 164 and an Industry Standard Architecture (ISA) bus 162 that connects to the System Interface Controller 106. Of course, other buses such as PCI, EISA and Microchannel may be used. The CPU 164 may be any conventional general purpose single-chip or multi-chip microprocessor such as a Pentium®, Pentium® Pro or Pentium® II processor available from Intel Corporation, a SPARC processor available from Sun Microsystems, a MIPS® processor available from Silicon Graphics, Inc., a Power PC® processor available from Motorola, or an ALPHA® processor available from Digital Equipment Corporation. In addition, the CPU 164 may be any conventional special purpose microprocessor such as a digital signal processor or a graphics processor.

The System Recorder 110 and Chassis controller 170, along with the NVRAM 112 that connects to the System Recorder 110, may be located on the backplane 152 of the server 100. The System Recorder 110 and Chassis controller 170 are typically the first microcontrollers to power up when server power is applied. The System Recorder 110, the Chassis controller 170 and the Remote Interface microcontroller 200 are the three microcontrollers that have a bias 5 volt power supplied to them. If main server power is off, an independent power supply source for the bias 5 volt power is provided by the RIB 104 (Figure 1). The Canister controllers 172-178 are not considered to be part of the backplane 152 because they are located on separate cards and are removable.

Each of the microcontrollers has a unique system identifier or address. The addresses are as follows in Table 1:

TABLE 1

Microcontroller	Address
System Interface controller 106	10
CPU A controller 166	03
CPU B controller 168	04
System Recorder 110	01
Chassis controller 170	02
Canister A controller 172	20
Canister B controller 174	21
Canister C controller 176	22
Canister D controller 178	23
Remote Interface controller 200	11

The microcontrollers may be Microchip Technologies, Inc. PIC processors in one embodiment, although other microcontrollers, such as an 8051 available from Intel, an 8751 available from Atmel, and a P80CL580 microprocessor available from Philips, could be utilized. The PIC16C74 (Chassis controller 170) and PIC16C65 (the other controllers) are members of the PIC16CXX family of CMOS, fully-static, EPROM-based 8-bit microcontrollers. The PIC controllers have 192 bytes of RAM, in addition to program memory, three timer/counters, two capture/compare/Pulse Width Modulation modules and two serial ports. The synchronous serial port is configured as a two-wire Inter-Integrated Circuit (I^2C) bus in one embodiment of the invention. The PIC controllers use a Harvard architecture in which program and data are accessed from separate memories. This improves bandwidth over traditional von Neumann architecture processors where program and data are fetched from the same memory. Separating program and data memory further allows instructions to be sized differently than the 8-bit wide data word. Instruction opcodes are 14-bit wide making

it possible to have all single word instructions. A 14-bit wide program memory access bus fetches a 14-bit instruction in a single cycle.

In one embodiment of the invention, the microcontrollers communicate through an I^2C serial bus, also referred to as a microcontroller bus 160. The document "The I^2C Bus and How to Use It" (Philips Semiconductor, 1992) is hereby incorporated by reference. The I^2C bus is a bidirectional two-wire bus that may operate at a 400 kbps. However, other bus structures and protocols could be employed in connection with this invention. For example, Apple Computer ADB, Universal Serial Bus, IEEE-1394 (Firewire), IEEE-488 (GPIB), RS-485, or Controller Area Network (CAN) could be utilized as the microcontroller bus. Control on the microcontroller bus is distributed. Each microcontroller can be a sender (a master) or a receiver (a slave) and each is interconnected by this bus. A microcontroller directly controls its own resources, and indirectly controls resources of other microcontrollers on the bus.

Here are some of the features of the I^2C-bus:

- Two bus lines are utilized: a serial data line (SDA) and a serial clock line (SCL).
- Each device connected to the bus is software addressable by a unique address and simple master/slave relationships exist at all times; masters can operate as master-transmitters or as master-receivers.
- The bus is a true multi-master bus including collision detection and arbitration to prevent data corruption if two or more masters simultaneously initiate data transfer.
- Serial, 8-bit oriented, bidirectional data transfers can be made at up to 400 kbit/second in the fast mode.

Two wires, serial data (SDA) and serial clock (SCL), carry information between the devices connected to the I^2C bus. Each device is recognized by a unique address and can operate as either a transmitter or receiver, depending on the function of the device. For example, a memory device connected to the I^2C bus could both receive and transmit data. In addition to transmitters and receivers, devices can also be considered as masters or slaves when performing data transfers (see Table 2). A master is the device which initiates a data transfer on the bus and generates the clock

signals to permit that transfer. At that time, any device addressed is considered a slave.

TABLE 2 Definition of I^2C-bus terminology

Term	Description
Transmitter	The device which sends the data to the bus
Receiver	The device which receives the data from the bus
Master	The device which initiates a transfer, generates clock signals and terminates a transfer
Slave	The device addressed by a master
Multi-master	More than one master can attempt to control the bus at the same time without corrupting the message
Arbitration	Procedure to ensure that, if more than one master simultaneously tries to control the bus, only one is allowed to do so and the message is not corrupted
Synchronization	Procedure to synchronize the clock signal of two or more devices

The I^2C-bus is a multi-master bus. This means that more than one device capable of controlling the bus can be connected to it. As masters are usually microcontrollers, consider the case of a data transfer between two microcontrollers connected to the I^2C-bus. This highlights the master-slave and receiver-transmitter relationships to be found on the I^2C-bus. It should be noted that these relationships are not permanent, but depend on the direction of data transfer at that time. The transfer of data would proceed as follows:

1) Suppose microcontroller A wants to send information to microcontroller B:
   - microcontroller A (master), addresses microcontroller B (slave);
   - microcontroller A (master-transmitter), sends data to microcontroller B (slave-receiver);
   - microcontroller A terminates the transfer.

2) If microcontroller A wants to receive information from microcontroller B:

- microcontroller A (master) addresses microcontroller B (slave);
- microcontroller A (master-receiver) receives data from microcontroller B (slave-transmitter);
- microcontroller A terminates the transfer.

Even in this situation, the master (microcontroller A) generates the timing and terminates the transfer.

The possibility of connecting more than one microcontroller to the I^2C-bus means that more than one master could try to initiate a data transfer at the same time. To avoid the chaos that might ensue from such an event, an arbitration procedure has been developed. This procedure relies on the wired-AND connection of all I^2C interfaces to the I^2C-bus.

If two or more masters try to put information onto the bus, the first to produce a 'one' when the other produces a 'zero' will lose the arbitration. The clock signals during arbitration are a synchronized combination of the clocks generated by the masters using the wired-AND connection to the SCL line.

Generation of clock signal on the I^2C-bus is the responsibility of master devices. Each master microcontroller generates its own clock signals when transferring data on the bus.

The command, diagnostic, monitoring and history functions of the microcontroller network 102 are accessed using a global network memory model in one embodiment. That is, any function may be queried simply by generating a network "read" request targeted at the function's known global network address. In the same fashion, a function may be exercised simply by "writing" to its global network address. Any microcontroller may initiate read/write activity by sending a message on the I^2C bus to the microcontroller responsible for the function (which can be determined from the known global address of the function). The network memory model includes typing information as part of the memory addressing information.

Using a network global memory model in one embodiment places relatively modest requirements for the I^2C message protocol.

- All messages conform to the I^2C message format including addressing and read/write indication.
- All I^2C messages use seven bit addressing.
- Any controller can originate (be a Master) or respond (be a Slave).
- All message transactions consist of I^2C "Combined format" messages. This is made up of two back-to-back I^2C simple messages with a repeated START condition between (which does not allow for re-arbitrating the bus). The first message is a Write (Master to Slave) and the second message is a Read (Slave to Master).
- Two types of transactions are used: Memory-Read and Memory-Write.
- Sub-Addressing formats vary depending on data type being used.

IV. REMOTE INTERFACE BOARD

Referring to Figure 3, the remote interface board (RIB) 104, previously shown in Figure 1, will now be described. The RIB is an interface between the microcontroller network 102 (Figure 1) of the server system 100 and an external client computer 122/124. The server system status and commands are passed through the RS232 connector port 204 at the client side of the RIB to the microcontroller network 102 on the server 100, controlled through the on-board PIC16C65 microcontroller 200. Signals in the microcontroller network 102 are transported by the microcontroller bus 160 (Figure 2). In one embodiment, the microcontroller bus 160 utilizes the I^2C bus protocol, previously described. The signals on the microcontroller bus 160 are received from the server 100 by the RIB 104 on the RJ-45 cable 103 and are translated by the PIC16C65 microcontroller 200 into an eight signal RS232 protocol. These RS232 signals are passed through a RS232 line transceiver 202, such as a LT1133A chip available from Linear Technology, with a baud rate capable of reaching the speed of 120 kbaud. A 25 pin D-Sub connector 204 connects to the other side of the line transceiver 202 and provides the point at which either the local client computer 122 or the server modem 126 makes a connection.

The two wire microcontroller bus 160 is brought in from the server 100 and passed to the microcontroller 200 using the RJ-45 cable 103 and RJ-45 connector 226.

A switch 228, such as a QS3126 switch available from Quick Logic, connects to the RJ-45 connector 226 and provides isolation for the data and clock bus signals internal and external to the RIB 104. If the RIB 104 and switch 228 have power, the switch 228 feeds the bus signals through to a microcontroller bus extender 230. Otherwise, if the switch 228 does not have power, the microcontroller bus 160 is isolated from the RIB 104. The bus extender 230 connects between the switch 228 and the microcontroller 200. The bus extender 230 is a buffer providing drive capability for the clock and data signals. In one embodiment, the bus extender 230 is a 82B715 chip available from Philips Semiconductor. Microcontroller 200 Port C, bit 3 is the clocking bit and Port C, bit 4 is the data line.

Communication with the server modem 126 is based on the RS232 protocol. The microcontroller 200 generates the receive and the transmit signals, where the signal levels are transposed to the RS232 levels by the LT1133A line transceiver 202. There are three transmit signals, RTS, SOUT and DTR, which are from Port A, bits 2, 3 and 4 of the microcontroller 200, whereas the five receive signals are from two ports, DCD, DSR from Port C, bits 1 and 0 and SIN, CTS and RI from Port A, bits 5, 0 and 1.

In one embodiment, the 25 pin RS232 pin connector 204 is used instead a 9 pin connector, since this type of connector is more common. All the extra pins are not connected except the pins 1 and 7, where pin 1 is chassis ground and pin 7 is a signal ground.

A static random access memory (SRAM) 208 connects to the microcontroller 200. In one embodiment, the SRAM 208 is a 32k x 8 MT5LC2568 that is available from Micron Technology. The SRAM 208 is also available from other memory manufacturers. An external address register 206, such as an ABT374, available from Texas Instruments is used for latching the higher addressing bits (A8-A14) of the address for the SRAM 208 so as to expand the address to fifteen bits. The SRAM 208 is used to store system status data, system log data from the NVRAM 112 (Figure 1), and other message data for transfer to the external interface port 204 or to a microcontroller on the microcontroller bus 160 (Figure 2).

Port D of the microcontroller 200 is the address port. Port B is the data bus for the bi-directional data interconnect. Port E is for the SRAM enable, output tristate and write control signals. The microcontroller 200 operates at a frequency of 12 MHz.

An Erasable Programmable Read Only Memory (EPROM) 212 is used for storing board serial number identification information for the RIB 104. The serial number memory 212 is signal powered, retaining the charge into a capacitor sourced through the data line. In one embodiment, the serial number memory 212 stores eight sixteen-byte serial/revision numbers (for maintaining the rework/revision history) and is a DS2502 chip available from Dallas Semiconductor. The programming of memory 212 is handled using a jumper applied through an external connector J1 210. The serial number memory 212 connects to the microcontroller 200 at Port C, bit 6 and to the external connector J1 210.

The RIB 104 may be powered through a 7.5 Volt/800mA supply unit that plugs into a connector J2 220. In one embodiment, the supply unit is 120 Volt AC to DC wall adapter. Connector J2 220 feeds a LT1376 high frequency switching regulator 222, available from Linear Technology, which regulates the power source. The regulated power output is used locally by the components on the RIB 104, and 300 mA are sourced to the microcontroller network 102 through a 300 mA fuse 224 and the RJ-45 connector 226. Thus, the output of the regulator 222 provides an alternative source for a bias-powered partition of the microcontroller network 102. The bias-powered partition includes the system recorder 110 (Figure 1), the NVRAM 112 and the Chassis controller 170 (Figure 2) which are resident on the server backplane 152.

V. REMOTE INTERFACE SERIAL PROTOCOL

The microcontroller network remote interface serial protocol communicates microcontroller network messages across a point-to-point serial link. This link is between the RIB controller 200 that is in communication with the Recovery Manager 130 at the remote client 122/124. This protocol encapsulates microcontroller network

messages in a transmission packet to provide error-free communication and link security.

In one embodiment, the remote interface serial protocol uses the concept of byte stuffing. This means that certain byte values in the data stream have a particular meaning. If that byte value is transmitted by the underlying application as data, it must be transmitted as a two-byte sequence.

The bytes that have a special meaning in this protocol are:

SOM 306	Start of a message
EOM 316	End of a message
SUB	The next byte in the data stream must be substituted before processing.
INT 320	Event Interrupt
Data 312	An entire microcontroller network message

As stated above, if any of these byte values occur as data in a message, a two-byte sequence must be substituted for that byte. The sequence is a byte with the value of SUB, followed by a type with the value of the original byte, which is incremented by one. For example, if a SUB byte occurs in a message, it is transmitted as a SUB followed by a byte that has a value of SUB+1.

Referring to Figure 4, the two types of messages 300 used by the remote interface serial protocol will be described.

1. Requests 302, which are sent by remote management (client) computers 122/124 (Figure 1) to the remote interface 104.
2. Responses 304, which are returned to the requester 122/124 by the remote interface 104.

The fields of the messages are defined as follows:

SOM 306	A special data byte value marking the start of a message.
EOM 316	A special data byte value marking the end of a message.
Seq.# 308	A one-byte sequence number, which is incremented on each request. It is stored in the response.

TYPE 310		One of the following types of requests:
	IDENTIFY	Requests the remote interface to send back identification information about the system to which it is connected. It also resets the next expected sequence number. Security authorization does not need to be established before the request is issued.
	SECURE	Establishes secure authorization on the serial link by checking password security data provided in the message with the microcontroller network password.
	UNSECURE	Clears security authorization on the link and attempts to disconnect it. This requires security authorization to have been previously established.
	MESSAGE	Passes the data portions of the message to the microcontroller network for execution. The response from the microcontroller network is sent back in the data portion of the response. This requires security authorization to have been previously established.
	POLL	Queries the status of the remote interface. This request is generally used to determine if an event is pending in the remote interface.
STATUS 318		One of the following response status values:
	OK	Everything relating to communication with the remote interface is successful.
	OK_EVENT	Everything relating to communication with the remote interface is successful. In addition, there is one or more events pending in the remote interface.
	SEQUENCE	The sequence number of the request is neither the current sequence number or retransmission request, nor the next expected sequence number or new request. Sequence numbers may be reset by an IDENTIFY request.

CHECK	The check byte in the request message is received incorrectly.
FORMAT	Something about the format of the message is incorrect. Most likely, the type field contains an invalid value.
SECURE	The message requires that security authorization be in effect, or, if the message has a TYPE value of SECURE, the security check failed.
Check 314	Indicates a message integrity check byte. Currently the value is 256 minus the sum of previous bytes in the message. For example, adding all bytes in the message up to and including the check byte should produce a result of zero (0).
INT 320	A special one-byte message sent by the remote interface when it detects the transition from no events pending to one or more events pending. This message can be used to trigger reading events from the remote interface. Events should be read until the return status changes form OK_EVENT to OK.

VI. RIB MICROCONTROLLER OPERATION

The remote interface is the bridge to link the microcontroller bus to the outside world via a RS232 serial port through which a client computer can be connected. A message from the remote client side via RS232 usually starts with the "Identify" command which identifies the system name. See the message format associated with Figure 4, above. The "Identify" command should be followed by the "Security" command with a password that is checked against the password stored in the NVRAM 112 (Figure 1). If the passwords match, the remote RS232 link is put in "secure mode" and the remote interface 104 (Figure 1) will now pass any "message" commands on to the microcontroller network bus 160 (Figure 2). Before the remote

application program disconnects the link, it should send the "Unsecure" command to take the RS232 link out of "secure mode".

Referring to Figures 5a and 5b, embodiments of the RIB microcontroller process 400 will be described. The process 400 is implemented as a computer program, termed firmware, written in PIC assembly language. The assembled machine code is stored in the microcontroller EPROM where each instruction is fetched for execution by the processor. The EPROM provides 4K x 14 program memory space, all on-chip. Program execution is using the internal memory. Of course, any of a variety of general purpose and special purpose processors could be used and the programming of the process 400 could be in high level code such as C or Java.

Beginning at an initialize PIC state 402, process 400 initializes the variables, stack pointer, and other structures of the RIB microcontroller 200 (Figure 3). Moving to state 404, a return point called "main" is identified in process 400. Proceeding to a decision state 406, process 400 determines if the RS232 port is transmitting data. If so, process 400 moves to state 408 to send a character (one byte) if there is data in the SRAM 208 to be sent out on the RS232 port 204. A process of receiving data via the RS232 port 204 is not shown herein. Receiving data via the port 204 is initiated by the use of an interrupt.

At the completion of state 408, or if decision state 406 evaluates to a false condition, process 400 proceeds to a Check Modem Status function 410 that is implemented as a modem dialing and answering state machine. Function 410 checks the status of the modem 126 for any possible activity. Function 410 will be further described in conjunction with Figure 6. Advancing to a decision state 412, process 400 determines if any server event is pending. Event types include, for example, CPU status change, power status change, canister status change, fan status change, temperature, and operating system timeout. If an event is pending, process 400 proceeds to state 414 and sends an event message to the client computer 122/124 via the RS232 port. If no event is pending, as determined at decision state 412, process 400 continues at a decision state 416. At decision state 416, process 400 checks to see if a RS232 remote message has been received from the client computer 122/124. If not, process 400 moves back to the "main" loop 404, as described above. One

reason that a message has not been received yet is that the modem is not yet transmitting.

If a message has been received, as determined at decision state 416, process moves to the appropriate state 420-426 to handle one of four command types: Identify, Secure, Unsecure, and Message. At state 420, process 400 performs the Identify command and identifies the system by responding with the system name retrieved from the System Recorder memory 112 (Figure 1).

At state 422, process 400 performs the Secure command and gets the password with the command and checks it against the password from the NVRAM 112 (Figure 1). If the passwords match, the access right is granted (opens secure mode), otherwise, reject the intent.

At state 424, process 400 performs the Unsecure command and releases the remote access right, i.e., closes secure mode. At the completion of states 420, 422 or 424, process 400 proceeds through off-page connector E 430 to state 438 (Figure 5b).

At state 426 on Figure 5b (through off-page connector D 418), process 400 performs the Message command and gets remote message data from the RIB SRAM 208 (Figure 3). Proceeding to a decision state 432, process 400 determines if this message command is for the remote interface 104. If it is, process 400 executes the internal remote interface function command, such as a Read Revision of the RIB command. If the message command is not for the remote interface, as determined by decision state 432, process 400 moves to state 436 and passes the message command to its destination (external to the remote interface) via the microcontroller bus. This facilitates communication with another microcontroller for a command to read or write information, for example.

At the completion of states 420, 422, 424, 434 or 436, process 400 advances to state 438 and stores the response data for the command into the SRAM 208 (Figure 3) to be sent back to the client computer 122/124. Moving to state 440, process 400 transmit the first byte of data back on the RS232 port 204 to the client computer 122/124. After the byte of data has been transmitted at state 440, process 400 moves back to the "main" loop 404 (on Figure 5a), as described above.

Referring to Figure 6, embodiments of the Check Modem Status function 410 will now be described. Function 410 is implemented as a modem dialing and answering state machine. Several terms useful for understanding of the modem dialing and answering state machine are listed in Table 3 below.

TABLE 3

Modem Term	Meaning
CTS	clear to send
DCD	data carrier detect
DSR	data set ready
DTR	data transfer ready
RTS	request to send
EOS	end of string
Protocol	indicates whether RS232 serial data uses the messaging protocol or whether the data is a string of bytes
Ring	modem is detecting an incoming ring signal from another modem
Local	a connection to a local client computer (no modem used)
Modem Mode	modem to modem connection
Modem Already Set	modem initialization string has already been sent and completed

State machine 410 includes nine states, states 470-486. State 470 denotes that the modem is disconnected, DTR and RTS are clear and the protocol is clear. Protocol is clear indicates that no message protocol processing is to occur for bytes on the RS232 link (because it would affect transmitting and receiving of modem control string bytes). The state machine 410 remains at the Modem Disconnect state 470 while CTS is clear OR there have been "n" dialing retries already OR there is no Ring OR DSR is clear. If DSR is set (active), the state machine 410 proceeds to a Local Modem state 486, wherein RTS and DTR are set. The state machine 410

remains at state 486 while DSR is set. Is DSR clears or if Local AND Modem Mode are both set, the state machine 410 returns to Modem Disconnect state 470.

The state machine 410 proceeds to Modem Soft Reset state 472 if a Call Out condition OR a Setup condition is achieved. Call Out is achieved if Modem Mode is set AND Modem Already Set is set AND CTS is set AND there have not been "n" dialing retries already. Setup is achieved if Modem Mode is set AND Modem Already Set is clear AND CTS is set. At Modem Soft Reset state 472, DTR is set and RTS is set. The state machine 410 remains at state 472 while Send String Done is clear, i.e., the modem command string is still being sent to the modem.

The state machine 410 proceeds to Modem Test state 474 when Send String Done is set. The state machine 410 remains at state 474 while Send String Done is clear. The state machine 410 proceeds to Modem Result Code state 476 when Send String Done is set. The state machine 410 remains at state 476 while Modem Result Status Done is clear, i.e., the results status of the modem test at state 474 is not yet available.

The state machine 410 returns to Modem Disconnect state 470 from state 476 if Results Status OK is clear, i.e., the results status is not OK. However, if Results Status OK is set, i.e., the results status is correct, the state machine 410 proceeds to a Modem Setup state 478, wherein Modem Already Set is set. The state machine 410 returns to Modem Disconnect state 470 from state 478 if there have been "n" dialing retries already. However, if there have not been "n" dialing retries already, the state machine 410 proceeds to a Modem Dialing state 480, wherein the modem is dialed.

The state machine 410 remains at state 480 while the previous EOS has not been reached AND two seconds have not passed. The state machine 410 returns to Modem Disconnect state 470 from state 480 if Dial OK is clear, i.e., dialing the modem was not successful. However, if Dial OK is set, i.e., dialing the modem was successful, the state machine 410 proceeds to a Modem Answering state 482. Another path to the Modem Answering state 482 is from the Modem Disconnect state 470 when a Ringing mode is achieved. Ringing mode is achieved if Modem Mode is set AND Modem Already Set is set AND CTS is set AND Ring is set. The state machine 410 remains at state 482 while DSR is clear OR DCD is clear. The state

machine 410 returns to Modem Disconnect State 470 from state 482 if DCD is clear and a timeout occurs, i.e., no DCD is set within a timeout period (nobody answers). The state machine 410 proceeds to Remote Modem state 484 when DSR is set AND DCD is set. The modem transfers message data while at this state. When DCD clears, the state machine 410 returns to Modem Disconnect state 470 from state 484 or otherwise remains at state 484.

While the above detailed description has shown, described, and pointed out the fundamental novel features of the invention as applied to various embodiments, it will be understood that various omissions and substitutions and changes in the form and details of the system illustrated may be made by those skilled in the art, without departing from the intent of the invention.